

Public Reporting of Quality Measures

What Are We Trying to Accomplish?

Robert M. Califf, MD, MACC, FAHA,* Eric D. Peterson, MD, MPH, FACC, FAHA†

Durham, North Carolina

More than a decade has passed since we first commented on both the promise and potential pitfalls of “cardiovascular scorecard medicine” (1,2). In this issue of the *Journal*, Resnic and Welt (3) report on their real-world experiences with public outcomes reporting for percutaneous coronary intervention (PCI) at Brigham and Women’s Hospital in Massachusetts. Their findings raise a number of fascinating issues that deserve careful consideration by cardiologists, administrators, researchers, policy makers, payers, and patients. Although some might see their work as an attempt to rationalize a mediocre performance rating, we believe their report serves a nobler purpose—it is the proverbial canary in the coal mine, warning of serious challenges inherent in relying upon in-hospital mortality as a meaningful measure of quality of care.

Public reporting of mortality outcomes in medicine has a long and complicated history, beginning, perhaps, with the ill-fated Dr. Earnest Codman. Dr. Codman, also from a prestigious Boston institution, daringly proposed to report his hospital’s mortality results, believing that this would attract informed consumers and would soon become standard practice for the country (4). Unfortunately for Dr. Codman, his hospital shortly went out of business and nobody followed his example.

Eighty years after the inglorious end of Dr. Codman’s experiment, believing that outcomes reporting would soon gain widespread acceptance, we predicted that:

The convergence of concern about costs of medical care, the availability of large amounts of clinical outcome data in computerized databases, and dramatic advances in the methods of assessing factors related to outcome have ushered in a new era of accountability for physicians, hospital and health care systems (1).

Although our proposal (like Dr. Codman’s) was a bit premature, a decade later, public reporting of hospital mortality data is now a reality for all U.S. cardiologists. In addition to the statewide activities discussed in this article,

the Center for Medicare and Medicaid Services (CMS) routinely reports 30-day hospital mortality rates for acute myocardial infarction and heart failure, and will soon add PCI mortality and heart failure readmission rates (5).

As we enter this new era of accountability, it is essential that the health services research community investigate both the intended and unintended consequences of these policies. Resnic and Welt (3) contribute to this evaluation by scrutinizing the public reporting process for PCI mortality in Massachusetts. Using their own hospital as an example, the investigators describe the challenges in benchmarking performance for acute PCI outcomes. They point out that present-day PCI mortality rates reflect relatively rare events that occur most commonly among patients who arrive at the catheterization laboratory with a multifactorial profile of extreme illness (e.g., in cardiogenic shock, with acute myocardial infarction, or with other major comorbid conditions). They also note that current risk-adjustment models often fail to fully capture these complexities, and they provide convincing evidence that public profiling efforts may have the unintended result of encouraging clinicians to avoid those high-risk patients most likely to receive the greatest benefit from the procedure.

Is the public release of PCI mortality data good or bad for our field? We might conclude that it is both, depending in part on the goal of such reporting. An oft-stated reason for public reporting programs is to inform consumers’ choices, so that they can select a quality provider and hospital for their PCI procedure. This rationale, however, falls flat for several reasons. First, mortality is a crude measure to apply to situations characterized by very low event rates, or for which the ability to predict the outcome depends upon factors that are poorly collected by current systems. As Resnic and Welt (3) note, PCI mortality occurred most often when it was performed under urgent or emergent conditions (and therefore, in circumstances in which consumer choice is moot). Under elective conditions, mortal complications related to PCI are exceedingly rare, making it unlikely that provider quality can be differentiated from chance background events. For example, an earlier study concluded that under plausible conditions, up to 90% of true poor PCI performers could be missed by a profiling system, and 60% to 70% of physicians identified as poor quality

From the *Duke Translational Medicine Institute and the †Duke Clinical Research Institute, Durham, North Carolina. Dr. Califf does not have specific industry relationships on this topic; a complete list of industry relationships can be found at <http://www.dcri.duke.edu/research/coi.jsp>. Dr. Peterson has received research support from Bristol-Myers Squibb/Sanofi, Schering-Plough, and Merck/Schering; additional disclosures can be found at <http://www.dcri.duke.edu/research/coi.jsp>.

Manuscript received October 14, 2008; accepted October 21, 2008.

providers might be falsely labeled, simply because of purely random variation (6). Additionally, public mortality data are 2 or more years out of date when reported. Because estimates for hospitals also tend to vary markedly by year, past outcomes information, like stock market warnings, may not reflect current or future performance.

On a more positive note, provider profiling efforts can in fact reinforce clinicians' and administrators' interest in the quality improvement (QI) process. Again, Resnic and Welt's paper (3) illustrates this effect. It is debatable whether Brigham and Women's Hospital would have assigned 3 full-time employees to support PCI clinical data collection and QI efforts had it not been for state mandates or the impending public release of performance data. Additionally, although the detailed review afforded to each PCI mortality event is laudable, such intensive QI was in part motivated by the desire to demonstrate flaws in the provider profiling system rather than purely to identify opportunities for preventing errors. Thus, although the public release of outcomes data may not necessarily serve the purpose of informing the public, it can nonetheless supply the strong external motivation needed to pursue internal QI activities.

Given that public reporting of death rates is likely here to stay, what can we do to make the best use of this information? Resnic and Welt (3) make several good suggestions, which we would like to highlight and augment: First, we are skeptical that any amount of tuning can resolve the difficult issue of predicting death after PCI with sufficient accuracy, or can overcome statistical challenges arising from the laws of small numbers. Although we do not support the public release of potentially misleading PCI mortality information, we do support state and national efforts that would mandate hospitals to collect and compare clinical information on PCI, such as that required in Massachusetts and offered voluntarily by the American College of Cardiology (ACC)–National Cardiovascular Data Registry (7). Armed with such data, we would further encourage all institutions to routinely review any deaths occurring during PCI in a manner similar to that outlined by Resnic and Welt (3), with the main goal of identifying opportunities for future care improvement.

In place of the public release of procedure-related outcomes, we would argue for better and more complete analysis of condition-specific outcomes (i.e., acute myocardial infarction and heart failure). For example, by examining all patients with MI, one can lessen concerns of provider "case selection creep," as both patients with and without a procedure will ultimately be included in the denominator. Yet in order to pursue such a course, we will need improved, comprehensive data, since current provider profiling efforts rely on administrative databases that are notoriously limited in terms of the completeness and accuracy needed for risk adjustment (8). Instead of implementing flawed profiling with inadequate administrative data, we again would suggest expanding participation in national clinical registries, such as those run by the ACC and the American Heart Association. In addition to providing incentives or man-

dates for participation, they could also provide support for auditing registries for completeness and accuracy and, ideally, for including longitudinal outcomes information.

We would also argue that any outcomes reporting system should be augmented with data on care process and appropriateness. Assuring that PCI patients receive optimal medical therapy may be as important as the procedure itself for long-term health outcomes, as demonstrated by recent trials (9). Thus, reporting on the routine use of these evidence-based secondary prevention therapies may increase their ubiquity, which might potentially save many more lives than reporting on small differences in the quality of the procedure itself could.

Similarly, as Resnic and Welt (3) acutely point out, selection of patients likely to benefit from the procedure is critical. Close monitoring of procedural appropriateness can help us avoid unnecessary risk to patients lacking good indications for a given procedure, while simultaneously reducing the associated financial costs. Procedural appropriateness criteria for PCI are being developed by the ACC to facilitate such measurement, and digital technology now permits remote oversight of angiograms to confirm anatomic findings and allow peer feedback. But again, external incentives must be in place before institutions will be willing to routinely collect and report these data.

Better education of the public and of provider communities on how to interpret and make use of this information is essential. We also need better methods for summarizing complex, multidimensional information on care structure, processes, and outcomes into composite measures that are meaningful and reflect their intended purpose and values (10).

Finally, professional regulations are needed to prevent inappropriate marketing of comparisons of quality data that are linked with advertisements promoting false or overstated impressions. Even in an era of public transparency and competition, the main role of quality assessment should be to promote QI, not increase market share.

In summary, Resnic and Welt (3) have done all of us in the cardiology community a favor by "singing" about the problems with the application of PCI mortality risk scores to assess comparative quality. As with any quality tool, the value of the measurement will be determined by the wisdom of those who use it.

Reprint requests and correspondence: Dr. Robert M. Califf, Vice Chancellor for Clinical Research, Director, Duke Translational Medicine Institute, Duke University Medical Center, Campus Box 3850, Durham, North Carolina 27710. E-mail: rob.califf@duke.edu

REFERENCES

1. Topol E, Califf R. Scorecard cardiovascular medicine. Its impact and future directions. *Ann Intern Med* 1994;120:65–70.
2. Califf RM, Jollis JG, Peterson ED. Operator specific outcomes: A call for professional responsibility. *Circulation* 1996;93:403–6.

3. Resnic FS, Welt FGP. The public health hazards of risk avoidance associated with public reporting of risk-adjusted outcomes in coronary intervention. *J Am Coll Cardiol* 2009;53:825–30.
4. Codman E. *A Study in Hospital Efficiency*. Boston, MA: Privately printed, 1916.
5. Center for Medicaid and Medicare Services. Technical specifications for proposed ACE demonstration quality measures requested in the acute care demonstration application. Available at: <http://www.hospitalcompare.hhs.gov/Hospital/Search/Welcome.asp>. Accessed October 13, 2008.
6. Thomas JW, Hofer TP. Accuracy of risk-adjusted mortality rate as a measure of hospital quality of care. *Med Care* 1999;37:83–92.
7. National Cardiovascular Data Registry. Available at: <http://www.ncdr.com/WebNCDR/Common/>. Accessed October 6, 2008.
8. Jollis JG, Ancukiewicz M, DeLong ER, Pryor DB, Muhlbaier LH, Mark DB. Discordance of databases designed for claims payment versus clinical information systems. Implications for outcomes research. *Ann Intern Med* 1993;119:844–50.
9. Boden WE, O'Rourke RA, Teo KK, et al. Optimal medical therapy with or without PCI for stable coronary disease. *N Engl J Med* 2007; 356:1503–16.
10. O'Brien SM, DeLong ER, Dokholyan RS, Edwards FH, Peterson ED. Exploring the behavior of hospital composite performance measures: an example from coronary artery bypass surgery. *Circulation* 2007;116:2969–75.

Key Words: percutaneous coronary intervention ■ quality assurance ■ cardiogenic shock.